

DATA BASE SYSTEM BASED ON VIDEO/AUDIO OBJECT

Publication number: JP2001167099 (A)

Publication date: 2001-06-22

Inventor(s): NAEUMURA MASAHIRO; IZUMI YOSHINORI; NISHI TAKASHI; FUKUDA ATSUSHI +

Applicant(s): JAPAN BROADCASTING CORP +

Classification:

- international: G06F12/00; G06F17/30; H04N5/91; H04N7/08; H04N7/081; G06F12/00; G06F17/30; H04N5/91; H04N7/08; H04N7/081; (IPC1-7): G06F17/30; G06F12/00; H04N5/91; H04N7/08

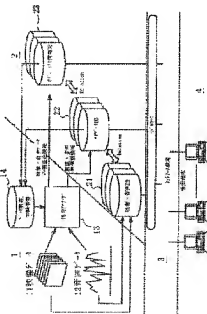
- European:

Application number: JP19990347094 19991207

Priority number(s): JP19990347094 19991207

Abstract of JP 2001167099 (A)

PROBLEM TO BE SOLVED: To provide a data base system capable of precisely handling video/ audio data with the object of video/audio as a basic unit.
SOLUTION: Concerning the data base system for managing video/audio data to change with the passage of time while using the object data of video/ audio and meta-data expressing the feature amounts of objects, this system is provided with a preprocessing part 1 for converting input video/audio data to be stored in a data base to a data format for the system, a data body part 2 for respectively relating the video/audio data converted by the preprocessing part, the object data and the meta-data and storing them in the data base, a network part 3 for providing an environment, in which the respective data stored in the data body part can be integrally handled while using middleware even in a distributed environment, and a user interface part 4 for retrieving required data from a user to the data base, the video/audio data, the object data and the meta-data are stored in respectively independent data bases and each of data is related by an identifier peculiar for these data.



Data supplied from the **espacenet** database — Worldwide

(51) Int.Cl. ⁷	識別記号	F I	テ-ロ-ド* (参考)	
G 0 6 F 17/30		G 0 6 F 12/00	5 1 2	5 B 0 7 5
12/00	5 1 2		5 4 7 D	5 B 0 8 2
	5 4 7	15/401	3 3 0 Z	5 C 0 6 3
H 0 4 N 5/91		15/40	3 4 0	5 C 0 6 3
7/08			3 7 0 D	

審査請求 未請求 請求項の数 2 O L (全 6 頁) 最終頁に続く

(21) 出願番号 特願平11-347094

(22) 出願日 平成11年12月7日 (1999.12.7)

(71) 出願人 000004352

日本放送協会
東京都渋谷区神南2丁目2番1号

(72) 発明者 苗村 昌秀

東京都世田谷区砧1丁目10番11号 日本放送協会 放送技術研究所内

(72) 発明者 和泉▲吉▼則

東京都世田谷区砧1丁目10番11号 日本放送協会 放送技術研究所内

(74) 代理人 100059258

弁理士 杉村 曉秀 (外2名)

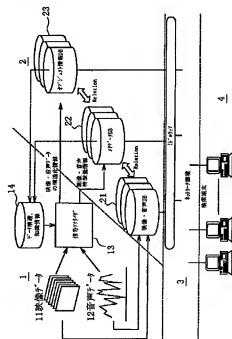
最終頁に続く

(54) 【発明の名称】 映像・音声オブジェクトを基本としたデータベース装置

(57) 【要約】

【課題】 映像・音声のオブジェクトを基本単位として、映像・音声データを細かく扱えるデータベース装置を提供する。

【解決手段】 時間的に変化する映像、音声データを映像・音声のオブジェクトデータおよびオブジェクトの特徴量を表すメタデータを用いて管理するデータベース装置において、当該装置が、データベースに格納する入力映像・音声データを前記装置用のデータフォーマットに変換する前処理部(1)と、該前処理部で変換された映像・音声データ、オブジェクトデータおよびメタデータをそれぞれ関連付けてデータベースに格納するデータ本体部(2)と、該データ本体部に格納されたそれぞれのデータをミドルウェアを用いて分散環境でも統合的に取り扱うことのできる環境を提供するネットワーク部(3)と、ユーザからのデータベースへの必要データの検索を行うユーザインターフェース部(4)とを具え、映像・音声データ、オブジェクトデータおよびメタデータをそれぞれ独立のデータベースに蓄え、それぞれのデータをそのデータ固有の識別子で関連付けるよう構成する。



【特許請求の範囲】

【請求項1】 時間的に変化する映像、音声データを映像・音声のオブジェクトデータおよびオブジェクトの特徴量を表すメタデータを用いて管理するデータベース装置において、

当該装置が、データベースに格納する入力映像・音声データを前記装置用のデータフォーマットに変換する前処理部と、該前処理部で変換された映像・音声データ、オブジェクトデータおよびメタデータをそれぞれ関連付けてデータベースに格納するデータベース本体部と、該データベース本体部に格納されたそれぞれのデータをミドルウェアを用いて分散環境でも統合的に取り扱うことのできる環境を提供するネットワーク部と、ユーザからのデータベースへの必要データの検索を行うユーザインターフェース部とを具え、映像・音声データ、オブジェクトデータおよびメタデータをそれぞれ独立のデータベースに蓄え、それぞれのデータをそのデータ固有の識別子で関連付けることを特徴とする映像・音声オブジェクトを基本としたデータベース装置。

【請求項2】 請求項1記載の装置において、前記データフォーマットが、映像・音声の原データであって番組を表す識別子および番組全体のインデックス情報を含む映像・音声データと、映像シーン毎のデータであって番組識別子、シーン識別子およびシーン全体のインデックス情報を含むシーン構造データと、抽出されたオブジェクト画像データの他に番組識別子、シーン識別子およびオブジェクト識別子を含むとともに前記オブジェクトを階層的に管理できるオブジェクト構造データおよびメタデータからなることを特徴とする映像・音声オブジェクトを基本としたデータベース装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 この発明はデータベースの構造に係り、特に映像・音声データベースを効率よく扱うことの可能な映像・音声オブジェクトを基本としたデータベース装置に関するものである。

【0002】

【従来の技術】 従来の映像・音声のデータベース技術としては、画面全体、音源全体からメタデータを抽出して、そのメタデータと映像・音声データを関連付けるのが一般的であった。このとき、メタデータが映像、音声を構成するオブジェクトに対して、関連付けられていなかったことで細かい制御が困難であった。

【0003】

【発明が解決しようとする課題】 従って本発明の目的は、前述の課題を解決し、従来の技術で取り扱うことができなかった映像・音声データの細かい制御の可能な、かつ動画における映像・音声データのオブジェクト単位での効果的な管理も可能な装置を提供せんとするものである。オブジェクト単位での管理によってより細かい映

像・音声データの管理が可能となる。

【0004】

【課題を解決するための手段】 前述の目的を達成するための、本発明映像・音声オブジェクトを基本としたデータベース装置は、時間的に変化する映像・音声データを映像・音声のオブジェクトデータおよびオブジェクトの特徴量を表すメタデータを用いて管理するデータベース装置において、当該装置が、データベースに格納する入力映像・音声データを前記装置用のデータフォーマットに変換する前処理部と、該前処理部で変換された映像・音声データ、オブジェクトデータおよびメタデータをそれぞれ関連付けてデータベースに格納するデータベース本体部と、該データベース本体部に格納されたそれぞれのデータをミドルウェアを用いて分散環境でも統合的に取り扱うことのできる環境を提供するネットワーク部と、ユーザからのデータベースへの必要データの検索を行うユーザインターフェース部とを具え、映像・音声データ、オブジェクトデータおよびメタデータをそれぞれ独立のデータベースに蓄え、それぞれのデータをそのデータ固有の識別子で関連付けることを特徴とするものである。

【0005】 また、本発明映像・音声オブジェクトを基本としたデータベース装置は、前記データフォーマットが、映像・音声の原データであって番組を表す識別子および番組全体のインデックス情報を含む映像・音声データと、映像シーン毎のデータであって番組識別子、シーン識別子およびシーン全体のインデックス情報を含むシーン構造データと、抽出されたオブジェクト画像データの他に番組識別子、シーン識別子およびオブジェクト識別子を含むとともに前記オブジェクトを階層的に管理できるオブジェクト構造データおよびメタデータからなることを特徴とするものである。

【0006】

【発明の実施の形態】 本発明によれば、映像・音声データ、オブジェクトデータおよびメタデータの関連付けは、それぞれのデータにオブジェクト毎の識別子を付加することによって行っているため、データベース内のオブジェクトのメタデータから映像・音声データにアクセスすることができるし、メタデータだけを操作してデータベース内のデータをオブジェクト単位に扱うこともできる。メタデータは映像・音声データより容量が小さいので操作性に優れており、オブジェクトを階層的な構造で管理することにより、任意の大きさのオブジェクトに対応することもできる。

【0007】 以下添付図面を参照し実施例により本発明の実施の形態を詳細に説明する。本発明に係るデータベース装置実施例構成図を図1に示す。本発明装置は前処理部1、データベース本体部2、ネットワーク部3およびユーザインターフェース部4の4つの部分より構成される。前処理部1は、データベースに格納する映像デー

タ11および音声データ12を本装置用のデータフォーマットに変換する部分である。映像・音声データ11、12は、信号アナライザ13によって構成オブジェクトに分解される。前処理部で重要なことは分解されたシーン、オブジェクトを識別するID（識別子）をつけることであり、分解作業自体はシーンチェンジ検出器、オブジェクト分解器などを使って簡単化することもできるし、マニュアル作業でシーンごと、映像、音声のオブジェクトごとにきめ細かく分解することもできる。このとき、信号アナライザ13は後述するデータベース本体部2から帰還されるデータを取り込んだデータ構造・知識情報メモリ14により制御されるし、シーン、オブジェクトの特徴を表すデータ（メタデータ）をも抽出し、シーン、オブジェクトを識別するIDと関連付けられる。メタデータとしては、オペレータがシーンごと、オブジェクトごとに付加するキーワード、オブジェクトの色情報、テキストや情報、動き情報などがある。

【0008】データベース本体部2は、映像・音声データ、オブジェクト情報データおよびメタデータそれぞれを関連付けてそれぞれのデータベース21、22、23に格納する部分である。図1では、映像・音声データは映像・音声DB（データベース）21に、オブジェクトに関する情報データはオブジェクト情報DB22に、メタデータはメタデータDB23にそれぞれ格納される。DBの管理は図2に示すデータフォーマットに基づいて行われる。

【0009】ここで、図2のデータフォーマットについて詳細な説明をする。映像・音声データ51とは映像・音声の原データを意味し、番組を表すIDと番組全体のインデックス情報を含む。シーン構造データ52とはシーン毎のデータであり、番組ID、シーンを表すシーンID、シーン全体のインデックス情報を含む。オブジェクト構造データ53とは抽出されたオブジェクトに関するデータのことで、オブジェクトそのものを表すオブジェクト画像データのほかに番組ID、シーンID、オブジェクトを表すオブジェクトIDを含む。オブジェクトの管理は粗いオブジェクトから細かいオブジェクトまで階層的にオブジェクトを扱えるように、図3に示す階層化構造を示す木構造でおこなわれる。

【0010】データベースではシーンごとに図3に示すような木構造を記述するデータを有する。これは、図3の場合、（#0、（#1、（#3、#4、#5、#12））、（#2、（#6、#7、#8、#9、#10、#11、#12）））のように、括弧の括弧の数がそのオブジェクトの階層位置を表すようなデータ

である。オブジェクトラベル#0は常に画面全体を表すものとする。なお、オブジェクトの画像は最下層のラベル画像で管理され、木構造との組み合わせで目的のオブジェクトを取り出せるようにしている。

【0011】図2にもどって、メタデータ54とはデータについて記述した特徴データのことであり、メタデータDBにはオブジェクト毎のメタデータが格納されている。メタデータの種類としてはテキストベースのキーワードと画像・音声ベースの色、テキストや、動き、周波数スペクトル情報などを含む。映像・音声データ、オブジェクトデータとの関連付けのため番組ID、シーンID、オブジェクトIDを含む。各データ間の関連付けは番組ID、シーンID、オブジェクトIDで行われる。これらのIDを用いて、音声・映像データをメタデータでオブジェクト単位に特定することができる。

【0012】たとえば、Oracle、IBM、Informixなどに代表されるオブジェクトリレーショナルDBで映像・音声DB、オブジェクト情報DB、メタDBを管理するときは、表1-3に示す複数のリレーショナルテーブルを用いて行う。表1-3で示されているテーブル群は、1つの番組に対応するものであり、番組IDごとにこのようなテーブルが存在する。テーブルの種類としてはシーンテーブル表1、オブジェクトテーブル表2、メタデータテーブル表3の3種類で映像・音声データ、オブジェクトデータ、メタデータを管理する。シーンテーブル表1ではシーンIDとそのシーンの時間位置を表すタイムコードよりなる。また、オブジェクトテーブル表2では分解されたオブジェクトを表すオブジェクトIDが、そのオブジェクトが現れるシーンのIDと語述するオブジェクト階層を示すIDの両方に関連付けられて管理されている。

【0013】

【表1】

シーン テーブル	
シーン番号	時間コード
#0	T0 T1
#0	T1-T2
#0	Tp-Tp+1

【0014】

【表2】

オブジェクト テーブル

シーン番号	オブジェクト階層	オブジェクトID	出現 T C
# 0	h 0	0 - 0	T0 - T1
...	h 1	0 - 1	T0(h1, 1) - T1(h1, 1)
...
# 0	h 1	0 - n 0	T0(h1, n0) - T1(h1, n0)
...
# p	# 0	p - 0	Tp - Tp + 1
...
# p	h 1	p - n p	Tp(h1, np) - Tp + 1(h1, np)
...

【0015】

【表3】

メタデータ テーブル

シー ン 番 号	オブ ジ ェ ク ト I D	メ タ デ ー タ 情 報				
		動き	形状	色	Text	音 特 性
# 0	0-0	M[0-0]	S[0-0]	C[0-0]	T[0-0]	A[0-0]
# 0	0-n0	M[0-n0]	S[0-n0]	C[0-n0]	T[0-n0]	A[0-n0]
# p	p-0	M[p-0]	S[p-0]	C[p-0]	T[p-0]	A[p-0]
# p	p-np	M[p-np]	S[p-np]	C[p-np]	T[p-np]	A[p-np]

画像データとの関連付け

シーン番号 11 — 1番目のシーンの原画
1番目のシーンのモニタ用画像
(代表フレーム、縮小圧縮画像)

$\text{objID} \quad m \quad n$ _____ m -番目のシーンの# n のオブジェクト輪郭
又はオブジェクト内での長方形

472:大陸層 41 i 番目の陽層での分割結果

【0016】それぞれオブジェクトは、シーン内での登場する時間を表すタイムコードデータと関連付けられる。このとき、オブジェクトは1シーンに複数回登場する場合もあるので(一度消えてからまた現れる場合など)、複数のタイムコードデータを有する。メタデータテーブルはオブジェクトの情報を表すメタデータを管理するもので、シーンID、オブジェクトIDと関連付けメタデータを管理する。メタデータの種類はオブジェクトの情報を表すものであればどんな種類のものでもよく、代表的なものとしてはオブジェクトの動きデータ、形状データ、色データ、キーワードおよび音のスペクトル情報などがあがる。

【0017】図1のネットワーク部3は、CORBAのようなミドルウェアを用いてオブジェクト情報DB、メタデータDB、映像・音声原DBを分散環境で統合的に

扱った部分である。具体的には、番組単位ごとにDBの名称をミドルウェアが提供するネーミングサービスに登録することによって、ネットワークに分散した環境でも、番組名を指定することによって、目的の番組を管理するデータベースにアクセスできる環境を提供する部分である。

【0018】同じく図1のユーザインターフェース部4は、ユーザからのデータベースへの必要データの検索を行う部分である。オブジェクトベースDBの場合、動画中のオブジェクトの指定が必要となる。以下に、オブジェクトを基本としたベースDBを用いた場合の、類似シーンの検索手順の一例を示す。

○番組の指定：番組IDの指定

ネットワーク部ミドルウェアが、ネーミングサービスに登録されたネットワークに接続されている複数のDBの

中から適切なDBを選択する。

○番組中から任意の1シーンをモニターする：シーンIDの指定

ネットワークを通してシーンの映像をユーザインターフェース部にダウンロードする。

○シーンの中から注目するオブジェクトと検索するメタデータを選択する：オブジェクト階層ID、オブジェクトID、メタデータの種類の特定

DBの中から選択されたオブジェクトに関連付けられたメタデータが連続される。

○検索を開始する

選択されたオブジェクトのメタデータと類似している（類似の定義は、ユーザで定義する）メタデータがメタデータDBより類似度の高い順番で検索される。

○検索されたメタデータに関連付けられているオブジェクト、シーンを類似度の高い順番で表示する。

このように、本発明の映像・音声オブジェクトを基本としたデータベース装置では、オブジェクト単位に映像・音声データが管理されているので、細かい検索が可能となる。

【0019】

【発明の効果】本発明によれば映像・音声データベース内の映像・音声データをオブジェクト単位で扱えるようになり、大量の映像・音声データの効果的な検索ができるようになる。応用例としては、サッカーなどのスポーツシーンから特定の選手が登場する場面を、その選手オブジェクトをキーとして検索することにより、高速検

索ができる。また、ユーザが興味あるオブジェクトだけを提示することにより、ネットワーク上での効率のよいQoS (Quality of Service) が実現でき。

【図面の簡単な説明】

【図1】 本発明を実現する装置のシステム構成図。

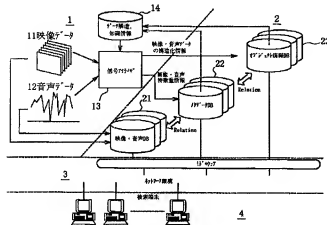
【図2】 データフォーマットを詳細に説明するための図。

【図3】 オブジェクトの階層化データ構造を示す図。

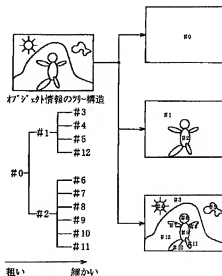
【符号の説明】

- 1 前処理部
- 2 データベース本体部
- 3 ネットワーク部
- 4 ユーザインタフェース部
- 11 映像データ
- 12 音声データ
- 13 信号アナライザ
- 14 データ構造、知識情報
- 21 映像・音声DB
- 22 メタデータDB
- 23 オブジェクト情報DB
- 51 映像・音声データ
- 52 シーン構造データ
- 53 オブジェクト構造データ
- 54 メタデータ
- 55 オブジェクト情報データ

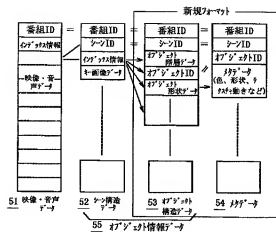
【図1】



【図3】



【図2】



フロントページの続き

(51)Int. Cl.⁷

H 0 4 N 7/081

識別記号

F I

H 0 4 N 5/91

ターマド (参考)

N

C

Z

7/08

(72)発明者 西 隆司

東京都世田谷区站1丁目10番11号 日本放

送協会 放送技術研究所内

F ターム(参考) 5B075 ND12 ND14 ND35 NK06 NR06

5B082 AA13 EA12 GA03

5C053 FA14 FA23 FA27 GB11 JA21

LA11

5C063 AC01 AC05

(72)発明者 福田 淳

東京都世田谷区站1丁目10番11号 日本放

送協会 放送技術研究所内